

TITLE OF THE INVENTION

IMAGE PERFECTION FOR VIRTUAL PRESENCE ARCHITECTURE (VPA)

[0001] This application claims priority of U.S. Provisional Patent Application Serial No. 60/452,275 filed March 4, 2003, which is hereby fully incorporated by reference.

FIELD OF THE INVENTION

[0002] This invention generally relates to the field of remote computer access. More specifically, an embodiment of the present invention relates to image perfection for virtual presence architectures.

BACKGROUND OF THE INVENTION

[0003] It is often the case that a host computer is located physically distant from its operator. Some products have been created to facilitate remote control of a computer using devices that remotely project the keyboard, video and mouse. These are typically called keyboard-video-mouse (KVM) devices. For example, a KVM switch enables a single keyboard, mouse and video display to be shared by multiple computers. A KVM device enables a keyboard, mouse and video display to be viewed remotely, with typically several hundred feet of separation. Remote Control Software enables a computer to "take over" a remote computer and use the local machine to provide keyboard and mouse input, and video output over a network. Additionally, there are specialized hardware components that interact with proprietary software to provide remote KVM functionality over a network.

[0004] However, each of the above approaches has some disadvantage. Software configuration of the host is one of the most difficult, in part, because it can differ significantly from machine to machine, for example, depending on the installed software and hardware. Also, any time additional hardware is added, other hardware issues may be introduced such as the need for platform certification, new drivers, and the like.

[0005] Consequently, a system is needed that is capable of remotely controlling a computer without interacting with the internal processing of that computer. Additionally, there are no currently available virtual presence devices that attach directly to the host computer which use only a slot, thus providing access to both power and ground from the host.

[0006] Additionally, there is no current system of perfecting the images sent and displayed using virtual presence architecture. Because images transferred from a host computer to a client computer can develop errors during compression, encoding or transmission, it is desirable to have software in the virtual presence architecture to remedy the errors and improve the image for display on the client computer.

BRIEF SUMMARY OF THE INVENTION

[0007] The present invention, which may be implemented utilizing a general-purpose digital computer, in certain embodiments of the present invention, includes novel methods and apparatus to provide efficient, effective, and/or flexible ability to automatically adjust a video signal without the need for user intervention.

[0008] In one embodiment of the present invention, there is an architecture that provides remote control of a host computer over existing Internet protocol (IP) network

infrastructure without requiring significant configuration changes, such as outside connections, to the remote host, but allows deployment with different levels of intrusiveness (e.g. depending on the requirements of the application).

[0009] In another embodiment of the invention, images transmitted from the host computer to the remote computer are improved so as to remove errors from the images displayed to the user of the client computer. In a further embodiment of the invention, the virtual presence architecture can automatically initiate and perform image improvement routines or can prompt a user to supply settings for the image perfection. In another embodiment of the invention, the virtual presence architecture uses phase locked loops to perform screen data comparisons and remove noise from images that are to be transmitted.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] Fig. 1 is a block diagram of an exemplary system into which virtual presence architecture may be implemented.

[0011] Fig. 2 is an exemplary block diagram of a virtual presence architecture.

[0012] Fig. 3 is a more detailed block diagram of a virtual presence architecture.

[0013] Fig. 4 illustrates an exemplary flow for the VPS and VPC video subsystems.

DETAILED DESCRIPTION OF THE INVENTION

[0014] Fig. 1 shows the basic format where the Video Presence Architecture (VPA) may be implemented. The computer system 100 comprises a central processor 102, a main memory 104, an input/output (I/O) controller 106, a keyboard 108, a pointing

device 110 (e.g. mouse, track ball, stylus, or the like), a display device 112, a mass storage 114 (e.g. hard disk, optical drive, or the like), and a network interface 118. Additional I/O devices, such as printing device 116, may be included in the computer system 100 as desired.

[0015] The system also comprises system bus 120, or similar architecture through which some or all of the components shown communicate with each other. Additionally, those with ordinary skill in the art will recognize that computer system 100 can include an IBM-compatible personal computer utilizing an Intel microprocessor, or any other type of computer. Additionally, instead of a single processor, two or more processors can be utilized to provide faster operations.

[0016] The network interface 118 provides communication capability with other computer systems on the same local network, on a different network connected via modems and the like to the present network, or to other computers across the Internet. In various embodiments, the network interface 118 can be implemented in Ethernet, Fast Ethernet, Gigabit Ethernet, wide-area network (WAN), leased line (such as T1, T3, optical carrier 3 (OC3), and the like), digital subscriber line (DSL and its varieties, such as high bit-rate DSL (HDSL), integrated services digital network DSL (IDSL) and the like), time division multiplexing (TDM), asynchronous transfer mode (ATM), satellite, cable modem, Universal Serial Bus (USB) and FireWire.

[0017] Fig. 2 illustrates an exemplary block diagram of a Virtual Presence Architecture (VPA) in accordance with an embodiment of the present invention.

[0018] Table 1 below provides a glossary of the terms used to describe the VPA architecture in accordance with some embodiments of the present invention (such as those discussed with respect to Figs. herein).

TERM	GLOSSARY
Capture	The process of digitizing and formatting data for processing.
Decode	Decode: the process of converting data encoded, e.g., by a virtual presence encoder for a device into a form suitable for transfer to that device.
Encode	The process of converting signals captured for a device into a form suitable for transfer to, e.g. a virtual presence decoder.
Host	The remote computer that is to be controlled from the local client.
NIC	Network interface connection, i.e., the device that provides network connectivity.
VPC	Virtual presence client; the subsystem that captures keyboard, mouse and other local device inputs for transmission to the VPS, and decodes the video display and other outputs from the VPS
VPP	Virtual presence protocol; the syntax and semantics of the messages exchanged by the VPS and the VPC. The VPP may be implemented on transmission control protocol (TCP) and user datagram protocol (UDP) over IP in an embodiment of the present invention
VPS	Virtual presence server; the subsystem that captures the hardware outputs of the host, encodes them for transmission to the VPC, and decodes the keyboard, mouse and other device inputs transmitted by the VPC.
Message Multiplexer	The entity that receives messages and tags them as being a particular type, then delivers them to be compressed and optionally encrypted.

TERM	GLOSSARY
Message Demultiplexer	The entity that takes decrypted and decompressed data from the stream and delivers it to the receiver registered to get that message type.
Frame Buffer	Memory where the digital image of the screen is stored; in an embodiment of the present invention, it consists of 16 bit pixels with 5 bits each for Red, Green and Blue intensity.
Tile	256 pixel area of the frame buffer treated as a unity by the video subsystem in accordance with an embodiment of the present invention.

Table 1 – Glossary of Terms

[0019] In Fig. 2, the VPA 200 includes a Virtual Presence Server (VPS) 204 co-located with the remote host 202 and a Virtual Presence Client (VPC) 208 at a location remote from the VPS. The host 202 interacts with the devices connected to the VPC (such as video display 210, keyboard 212, mouse 214, and other device 216) as if they were connected directly to host 202. In one embodiment of the present invention, an advantage of this approach is the flexibility in the design and deployment of the VPS 204.

[0020] Fig. 2 further demonstrates that keyboard 212, mouse 214, other device 216 send their respective signals to the VPC 208. VPC 208 captures the hardware outputs of the host and encodes them for transmission to the VPS. The transmission to the VPS can take place over IP Network 206, which is connected to host computer 202. Following transmission, the signals arrive in VPS 204, which decodes the keyboard, mouse and other device inputs transmitted by the VPC. These inputs are then sent to the host computer, where the input commands are executed. Following the execution of the keyboard, mouse and other device commands, host 202 sends a hardware output in the

form of a video signal displaying changes resulting from the input commands and a signal for the other device 216. The VPS 204 captures the hardware outputs and encodes them for transmission to the VPC 208 over IP Network 206. VPC 208 then decodes the video and other device outputs from the VPS and transmits them to either video display 210 or other device 216.

[0021] Fig. 3 illustrates a more detailed block diagram of a VPA in accordance with another embodiment of the present invention. Here, VPC 305 accepts signals from keyboard 348, mouse 350, and other device 352. These signals are then input to Keyboard Logic 354, Mouse Logic 362, and Other Device Logic 370, respectively. Inside each of the logic devices, the respective signals are captured at steps 356, 364, and 372, respectively, and are digitized and formatted for processing at steps 358, 366, and 374, respectively. After processing, the signals are encoded at steps 360, 368, and 376, respectively, by converting the captured signals for each device into forms suitable for transfer to a decoder, such as a Virtual Presence decoder. After the signals are encoded, they are sent to multiplexer 380, which combines the keyboard, mouse and other device signals in preparation for transmission to the VPS 304. However, before transmission, the signals can optionally be compressed in step 382 and/or encrypted in step 384. Then the signals are transported in 386 via IP network 344 to the VPS 304.

[0022] Once in VPS 304, the signals are decrypted and decompressed in items 332 and 334, respectively, if required. The input signals are then demultiplexed in 336 in order to separate the signals for decoding in items 338, 340, and 342. Then the keyboard, mouse and other device signals are sent to the host 302, where the commands are executed internally. Following the execution of the keyboard, mouse and other device

inputs, two hardware output signals are transmitted back to VPS 302, the video output signal and the other device output signal. The video output signal enters Video Logic element 306, which captures, compares, analyzes and encodes the output in steps 308-314, respectively. The other device output signal is sent to Other Device Logic element 316, where it is captured, processed and encoded in steps 318-322, respectively. The encoded video and other device outputs are then multiplexed in step 324, and can optionally be compressed and/or encrypted in steps 326 and 328, respectively.

[0023] The multiplexed output signal is then transported in step 330 over IP Network 344 to the VPC 305. Once the output signal is back in the VPC, it is decrypted and decompressed, if need be, in steps 390 and 392, respectively. The output signal is then demultiplexed into separate video and other device signals in step 394. Following that, the two signals are decoded in steps 396 and 398, and then sent to video display 346 and other device 352, where the outputs are displayed to the remote user. For example, the video output signal of host 302 is displayed on video display 346, and the other device output signal is executed on other device 352.

[0024] In another embodiment of the present invention, the devices in the VPA can be characterized by their data flow requirements. For example, the video logic system 306 on the VPS captures video frames, does delta analysis and encodes the stream for the VPC to decode and display. This does not require any return information in accordance with an embodiment of the present invention. Similarly, the mouse and keyboard subsystems may simply transmit the stream from their corresponding devices on the VPC for transmission to the VPS. On the other hand, special devices such as USB

may require bi-directional transfer, which are treated as independent directional flows by the architecture.

[0025] In a further embodiment of the present invention, the VPS captures video and transmits it to the VPC. For example, the VPS receives the mouse and keyboard data streams from the VPC and decodes them into signals for the Host. The VPS manages input and output data streams for other devices and simulates the local interactions necessary to provide remote functionality.

[0026] In accordance with another embodiment of the present invention, the keyboard and mouse may both be simple byte streams. Therefore, there would be little processing necessary to decode the streams. However, there is significant processing to maintain synchronization and duplicate the semantics and timing of the streams so that the Host can properly maintain its states as if the devices were directly connected.

[0027] More specifically, in an embodiment of the present invention, the VPS keyboard subsystem relays the byte stream from the remote keyboard to the Host without any additional processing. In a further embodiment, the VPS mouse subsystem relays the byte stream from the remote mouse to the Host. This byte stream may include “delta” messages (e.g. indicating change), which are interpreted by the Host relative to the current position of the cursor. Due to timing and other issues, the relative position of the cursor can get out of sync. Consequently, special processing in both the VPS and VPC can be used to mitigate this problem.

[0028] Fig. 4 illustrates an exemplary flow for the VPS and VPC video subsystems in accordance with an embodiment of the present invention. Since video is often the most data intensive part of the Virtual Presence system, the most significant

processing occurs in this component. VPS Video subsystem 400 captures the red, green and blue (RGB) video signals outputted from a host computer in step 402. The RGB signals are then transmitted to Current Frame Buffer 404. The illustrated video subsystem may be implemented in accordance with two characteristics of a computer's video display (such as the system discussed with respect to Fig. 1). Because screens may be primarily one color and because the screen typically only changes in local areas, leaving most of the display the same, the Video subsystem can take advantage of these characteristics to provide significant reductions in the required data.

[0029] In an embodiment of the present invention, the VPS video logic may specifically benefit from the creation of custom hardware to support the process. In another embodiment of the present invention, a field-programmable gate-array (FPGA) may be utilized to implement the logic in hardware. Further information regarding an FPGA apparatus for a VPA is later described in detail.

[0030] For example, in one embodiment, the video may be first captured into one of two frame buffers that alternate between being the current frame buffer 404 and the last frame buffer 406. In the present embodiment, the frame buffer is divided into "tiles" of 256 pixels. The Monochrome detection logic 408 analyzes each tile to see if its pixels are within a specified difference in color. If they are, then the Monochrome Map 410 corresponding to that tile receives a 1; otherwise, it receives a 0. The Difference Detection logic 412 compares each pixel in the Current frame buffer 404 with the corresponding pixel in the Last frame buffer 406. If more than a specified number of pixels have changed, then the bit corresponding to this tile is set to one in the Difference Map 414; otherwise it is set to zero.

[0031] In accordance with another embodiment of the present invention, the video encoder 416 then processes the two maps minimizing the data transmitted to indicate which tiles are changed, and sending a “raw” tile or a “monochrome” tile or a “no change” tile, and using, for example, run-length encoding to eliminate duplicates. The encoded stream is then passed to the message delivery subsystem 418 for optional compression and encryption, and then transmission to the VPC 420.

[0032] In a further embodiment of the present invention, the VPC 420 captures keyboard and mouse data streams, encodes them, and transmits the streams to the VPS 400. The VPC 420 later receives an encoded video stream, decodes it in step 422, and then processes the stream to remove encoding artifacts in step 424. The VPC then transfers the image to its own display, mapping the pixel image as needed. In particular, because the mouse is used as a pointing device and its motion is translated to a cursor on the video image, special processing may be utilized to keep the VPC cursor synchronized with the Host cursor.

[0033] Alternatively, other devices may be remotely connected to the host using a similar architecture. For example, a USB device, which provides a serial connection to deliver a stream of bytes between two entities, may be remotely connected to the host. USB devices have certain timing and signaling characteristics that are required for its function. Further, because USB devices are bi-directional, a complete encode and decode subsystem may be implemented for both VPS and VPC.

[0034] Moreover, the VPS may implement the logic necessary to emulate the USB device for the Host. Additionally, the VPC may implement the logic necessary to emulate the Host for the USB device. This process may require buffering of the byte

stream on both ends and emulating the timing characteristics required. This may also require special processing similar to the video subsystem depending on the particular device (such as that discussed with respect to Fig. 4). In particular, new digital display devices can be used to replace more traditional cathode ray tubes (CRTs) in many applications, and can be connected using USB technology.

[0035] In another embodiment of the present invention, the VPC encodes the byte stream from the local keyboard and delivers it to the message subsystem, which in turn optionally compresses and encrypts the stream. The stream is then transmitted to the VPS. Keyboard processing is envisioned to be a simple direct transfer with no feedback between the VPS and VPC in accordance with an embodiment of the present invention.

[0036] In further embodiment of the present invention, the VPC encodes the byte stream from the local mouse and delivers it to the message subsystem, which in turn optionally compresses and encrypts the stream, and then delivers the stream to the VPS. The encoding consists of aggregating mouse-move messages and transmitting them. Additional processing may be performed by the mouse subsystem to keep the cursors synchronized.

[0037] In another embodiment of the present invention, the VPC receives an encoded video stream from the VPS. The VPC decodes the stream into a working buffer, which it then processes to remove artifacts of the encoding algorithm used. Then the working buffer is transmitted to the actual display buffer on the VPC, which the video hardware displays on the local display device.

[0038] With respect to adjustment of the parameters in the ADC and capture path, in one embodiment of the present invention, the VPS program may stop and ask the user

for help adjusting these values or allow the user to enter a configuration screen to adjust them. In accordance with an embodiment of the present invention, the adjustments may be made automatically. In one embodiment of the present invention, on each new screen resolution that is received by the VPS, the VPS adjusts the borders of the screen. It performs this by setting the capture engine to move the screen down and to the right. Then it examines the memory to search for the black borders. If no borders are found that are close to where the Video Electronics Standards Association (VESA) specification says they should be (for example, with respect to an IBM-compatible PC), then most likely there is a large amount of real black space on the screen and the VESA values are loaded.

[0039] In a further embodiment of the present invention, each time a new screen resolution is detected, the VPS will enter a phase locked loop (PLL) adjustment cycle. A wide range of PLL values are tried and an algorithm detects the best one. In one embodiment of the present invention, each attempt includes capturing two screens together (e.g., within 25 milliseconds) and comparing them. The results of the compare are stored in a table. If no good match is found, the Tile Color Sensitivity is adjusted and a complete set of values are tried again. This is done up to several times. If no good set of values is found, then the values may be reverted back to a set of original settings, such as the Video Electronic Standards Association (VESA) standards, and tried again later. One reason that no PLL lock may be found is if the video screen is significantly changing, such as by screen saver.

[0040] In another embodiment of the present invention, a different algorithm can be used on a subsequent try. For example, on each screen update that is sent to the client,

the number of tiles that changed and the area where the changes happened are sent to an automatic adjustment module. This module can examine the area and the number of changes to determine if the changes are due to a bad PLL lock or valid data changing. If the module determines that there is a bad lock, it may slowly adjust the PLL parameters and see what difference the slow adjustment makes and provide for necessary adjustment later.

[0041] When the screen shots are taken, it is possible that there can be randomly scattered tiles over the entire image. These changes will likely be noise. Additionally, if, one third of the screen has changed, for example, but that change takes place over the entire screen, then it is most likely noise and should be filtered out. Further, if there are only smaller changes on the screen, then it is a real change in the display data and should not be interpreted as noise.

[0042] Therefore, the screen capture and compare algorithm discussed previously can be implemented to generate statistics. If the algorithm consistently sees that a certain amount (above a threshold value) of the screen is changing, it will interpret that change as noise and filter it out. However, if only a small portion of the screen is changing, or if a certain amount of the tiles are changing at the same time, the changes will be interpreted as real video data changes, and not as noise. In these situations, the algorithm will not adjust the video parameters.

[0043] It is envisioned that the architecture discussed herein may be implemented in many different ways. In various embodiments of the present invention, the Virtual Presence Architecture may be implemented utilizing one or more different techniques. For example, a heavily pipelined application specific integrated circuit (ASIC) or FPGA

to create the Tile Map and the Monochrome Map may be used. Also, when compressing and sending large data blocks, they may be split up so they overlap (for example: compress a little, send a little). Further, DIB Section application programming interfaces (API's) on Windows, or DirectX may be used. Additionally, to enhance compression, the extents of the changed area on the display can be detected and only info for that area may be sent. Also, the client may start the request for a next update area before it processes a current area, or the server may automatically prepare the next update area. Further, if there is more than one Monochrome or No Change tile, they may be stacked together and sent as one count. Speed can also be increased by overlapping as many operations as possible that can happen in parallel and, for example, blending the edges with a surrounding area when painting the monochrome tile on the client. Further, for slower links such as Dial-up or DSL, the packet turn around time can be relatively long, so one can modify any transport used to send long streams of packets and not spend time waiting for acknowledgements. Also, a compression function can be picked that is balanced in time with the transport time (for example, one may avoid spending more time compressing than the bandwidth of the transport may easily handle). Also, the client code can be tuned to the native OS and CPU for best performance. Finally, for very slow transports, extra time can be spent to break up tiles into subsections, and reduce data (e.g., blend groups or pixels into one, or reduce to 8-bit color instead of 32-bit color, and the like).

[0044] The foregoing description has been directed to specific exemplary embodiments of the present invention. It will be apparent to those with ordinary skill in the art that modifications may be made to the described embodiments of the present

invention, with the attainment of all or some of the advantages. For example, the techniques of the present invention may be utilized for provision of remote situations, gaming and the like. Therefore, it is the object of the appended claims to cover all such variations and modifications as come within the spirit and scope of the invention.